Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	1	("0597817").PN.	US-PGPUB; USPAT	OR	OFF	2005/08/30 07:24
L2	1	("5978817").PN.	US-PGPUB; USPAT	OR	OFF	2005/08/30 07:32
L3	1	("6009459").PN.	US-PGPUB; USPAT	OR	OFF	2005/08/30 07:26
L4	1	("5987508").PN.	US-PGPUB; USPAT	OR	OFF	2005/08/30 07:32
L7	9	(US-20020038348-$ or US-20020033844-$).did. or (US-6662230-$ or US-6516337-$ or US-6434536-$ or US-6282567-$ or US-6269361-$ or US-5752022-$ or US-6516311-$).did.	US-PGPUB; USPAT	OR	ON	2005/08/30 08:41
L8	8	7 & ((report status) sanme spider)	US-PGPUB; USPAT	OR	ON	2005/08/30 08:42
L9	1	7 & ((report status) same spider)	US-PGPUB; USPAT	OR	ON	2005/08/30 08:43
L10	6	7 & (report status)	US-PGPUB; USPAT	OR	ON	2005/08/30 08:53
L11	1	7 & cookies	US-PGPUB; USPAT	OR	ON	2005/08/30 08:58
L12	6	7 & track$3	US-PGPUB; USPAT	OR	ON	2005/08/30 08:59
L13	0	7 & (track$3 same traffic)	US-PGPUB; USPAT	OR	ON	2005/08/30 08:59
L14	6	7 & traffic	US-PGPUB; USPAT	OR	ON	2005/08/30 09:03
L15	4	7 & (web same traffic)	US-PGPUB; USPAT	OR	ON	2005/08/30 09:55
L16	7	7 & (email e-mail mail pop)	US-PGPUB; USPAT	OR	ON	2005/08/30 10:36
L17	11	virtual same web same page$1 same (spider crawler robot)	US-PGPUB; USPAT	OR	ON	2005/08/30 10:37
S1	8	("20020033844" "6662230" "5752022" "6282567" "6516337" "20020038348" "6434536" "6269361").pn.	US-PGPUB; USPAT	OR	ON	2005/08/29 17:17
S2	9	S1 or ("6516311").pn.	US-PGPUB; USPAT	OR	ON	2005/08/29 18:10
S3	2	"virtual web page" same spider	US-PGPUB; USPAT	OR	ON	2005/08/29 18:14
S4	11	"Wirth, John JR "	US-PGPUB; USPAT	OR	ON	2005/08/29 18:13

S5	1	S4 & "virtual web page"	US-PGPUB; USPAT	OR	ON	2005/08/29 18:13
S6	1	S4 & virtual	US-PGPUB; USPAT	OR	ON	2005/08/29 18:13
S7	1	"Nelson, Todd C"	US-PGPUB; USPAT	OR	ON	2005/08/29 18:14
S8	23	"virtual web page"	US-PGPUB; USPAT	OR	ON	2005/08/29 18:14

Terms used **web spider stopper** Found **2,770** of **160,457**

Sort results by: relevance

Display results: expanded form

Save results to a Binder

Search Tips

Open results in a new window

Try an Advanced Search

Try this search in The ACM Guide

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next

Best 200 shown Relevance scale

1 Personalized spiders for web search and analysis

Michael Chau, Daniel Zeng, Hinchun Chen

January 2001 **Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries**

Full text available: pdf(672.04 KB) Additional Information: full citation, abstract, references, citings, index terms

Searching for useful information on the World Wide Web has become incr easingly difficult. While Internet search engines have been helping people to search on the web, low recall rate and outdated indexes have become more and more problematic as the web grows. In addition, search tools usually present to the user only a list of search results, failing to provide further personalized analysis which could help users identify useful information and comprehend these results. To alleviate these ...

Keywords: information retrieval, internet searching and browsing, internet spider, noun-phrasing, personalization, self-organizing map

2 Teaching key topics in computer science and information systems through a web search engine project

Michael Chau, Zan Huang, Hsinchun Chen

September 2003 **Journal on Educational Resources in Computing (JERIC)**, Volume 3 Issue 3

Full text available: pdf(169.15 KB) Additional Information: full citation, abstract, references, index terms

Advances in computer and Internet technologies have made it more and more important for information technology professionals to acquire experience in a variety of aspects, including new technologies, system integration, database administration, and project management. To provide students with a chance to acquire such skills, we designed a project called "Build Your Search Engine in 90 Days," in which students were required to build a domain-specific Web search engine in a semester. In this pa ...

Keywords: education, indexing, web computing, web search engine, web spiders

3 On network-aware clustering of Web clients

Balachander Krishnamurthy, Jia Wang

August 2000 **ACM SIGCOMM Computer Communication Review , Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication**, Volume 30 Issue 4

Full text available: pdf(568.99 KB) Additional Information: full citation, abstract, references, citings, index terms

Being able to identify the groups of clients that are responsible for a significant portion of a

Web site's requests can be helpful to both the Web site and the clients. In a Web application, it is beneficial to move content closer to groups of clients that are responsible for large subsets of requests to an origin server. We introduce clusters---a grouping of clients that are close together topologically and likely to be under common administrative control. We identify clu ...

4 A Web Crawler in Perl
Mike Thomas
August 1997 **Linux Journal**
Full text available: html(14.82 KB) Additional Information: full citation, index terms

5 Session 1D: self-organizing systems: How social spiders inspired an approach To region detection
Christine Bourjot, Vincent Chevrier, Vincent Thomas
July 2002 **Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1**
Full text available: pdf(666.51 KB) Additional Information: full citation, abstract, references, index terms

Reactive problem solving is a way to propose systems composed of simple interacting agents that collectively solve problems outside the scope of individual perceptions. In this domain, natural social systems are sources of inspiration for simple mechanisms.This article presents an approach to region detection inspired by social spiders. Based on a behavioral model determined by the simulation of collective weaving, we describe how we transposed it to obtain an approach for region detection in gr ...

Keywords: biological inspiration, reactive multi-agent system, region detection

6 Session 7B: Crawling on web graphs
Colin Cooper, Alan Frieze
May 2002 **Proceedings of the thiry-fourth annual ACM symposium on Theory of computing**
Full text available: pdf(220.93 KB) Additional Information: full citation, references, citings, index terms

7 Politics on the web: making political candidates flies instead of spiders
Nick Hopper
September 1996 **ACM SIGCAS Computers and Society**, Volume 26 Issue 3
Full text available: pdf(451.69 KB) Additional Information: full citation, references

8 Information retrieval on the web
Mei Kobayashi, Koichi Takeda
June 2000 **ACM Computing Surveys (CSUR)**, Volume 32 Issue 2
Full text available: pdf(213.89 KB) Additional Information: full citation, abstract, references, citings, index terms

In this paper we review studies of the growth of the Internet and technologies that are useful for information search and retrieval on the Web. We present data on the Internet from several different sources, e.g., current as well as projected number of users, hosts, and Web sites. Although numerical figures vary, overall trends cited by the sources are consistent and point to exponential growth in the past and in the coming decade. Hence it is not surprising that about 85% of Internet user ...

Keywords: Internet, World Wide Web, clustering, indexing, information retrieval, knowledge management, search engine

9 Discovering parallel text from the World Wide Web
Jisong Chen, Rowena Chau, Chung-Hsing Yeh
January 2004 **Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation - Volume 32 CRPIT '04**
Full text available: pdf(249.63 KB) Additional Information: full citation, abstract, references

Parallel corpus is a rich linguistic resource for various multilingual text management tasks, including cross-lingual text retrieval, multilingual computational linguistics and multilingual text mining. Constructing a parallel corpus requires effective alignment of parallel documents. In this paper, we develop a *parallel page identification system* for identifying and aligning parallel documents from the World Wide Web. The system crawls the Web to fetch potentially parallel multilingual W ...

10 Literate programming
Christopher J. Van Wyk
September 1989 **Communications of the ACM**, Volume 32 Issue 9
Full text available: pdf(586.07 KB) Additional Information: full citation, references, citings, index terms

11 Novel search environments: Comparison of two approaches to building a vertical search tool: a case study in the nanotechnology domain
Michael Chau, Hsinchun Chen, Jialun Qin, Yilu Zhou, Yi Qin, Wai-Ki Sung, Daniel McDonald
July 2002 **Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries**
Full text available: pdf(859.29 KB) Additional Information: full citation, abstract, references, citings, index terms

As the Web has been growing exponentially, it has become increasingly difficult to search for desired information. In recent years, many domain-specific (vertical) search tools have been developed to serve the information needs of specific fields. This paper describes two approaches to building a domain-specific search tool. We report our experience in building two different tools in the nanotechnology domain -- (1) a server-side search engine, and (2) a client-side search agent. The designs of ...

Keywords: indexing, information retrieval, internet searching and browsing, internet spider, noun-phrasing, personalization, post-retrieval analysis, self-organizing map, summarization, vertical search engine, web search engine

12 Web clustering and usage mining: Evaluation of web usage mining approaches for user's next request prediction
Mathias Géry, Hatem Haddad
November 2003 **Proceedings of the 5th ACM international workshop on Web information and data management**
Full text available: pdf(314.69 KB) Additional Information: full citation, abstract, references, citings, index terms

Analysis of Web server logs is one of the important challenge to provide Web intelligent services.In this paper, we describe a framework for a recommender system that predicts the user's next requests based on their behaviour discovered from Web Logs data. We compare results from three usage mining approaches: association rules, sequential rules and generalised sequential rules. We use two selection rules criteria: highest confidence and last-subsequence. Experiments are performed on three colle ...

Keywords: association rules, evaluation, frequent generalised sequences, frequent sequences, web usage mining

13 Intelligent agents for retrieving chinese Web financial news
Christopher C. Yang, Alan Chung
December 2000 **Proceedings of the twenty first international conference on Information systems**
Full text available: pdf(449.26 KB) Additional Information: full citation, references, citings, index terms

14 Learning classifiers: Using urls and table layout for web classification tasks
L. K. Shih, D. R. Karger
May 2004 **Proceedings of the 13th international conference on World Wide Web**
Full text available: pdf(357.43 KB) Additional Information: full citation, abstract, references, index terms

We propose new features and algorithms for automating Web-page classification tasks such as content recommendation and ad blocking. We show that the automated classification of Web pages can be much improved if, instead of looking at their textual content, we consider each links's URL and the visual placement of those links on a referring page. These features are unusual: rather than being scalar measurements like word counts they are *tree structured*---describing the position of the item ...

Keywords: classification, news recommendation, tree structures, web applications

15 Corpus Linguistics: Mining the web to create minority language corpora
Rayid Ghani, Rosie Jones, Dunja Mladenić
October 2001 **Proceedings of the tenth international conference on Information and knowledge management**
Full text available: pdf(1.41 MB) Additional Information: full citation, abstract, references, citings, index terms

The Web is a valuable source of language specific resources but the process of collecting, organizing and utilizing these resources is difficult. We describe CorpusBuilder, an approach for automatically generating Web-search queries for collecting documents in a minority language. It differs from pseudo-relevance feedback in that retrieved documents are labeled by an automatic language classifier as relevant or irrelevant, and this feedback is used to generate new queries. We experiment with var ...

16 Full Papers: Exposing document context in the personal web
David Wolber, Michael Kepe, Igor Ranitovic
January 2002 **Proceedings of the 7th international conference on Intelligent user interfaces**
Full text available: pdf(295.10 KB) Additional Information: full citation, abstract, references, citings, index terms

Reconnaissance agents show context by displaying documents with similar content to the one(s) the user currently has open. *Research paper search engines* show context by displaying documents that cite or are cited by the currently open document(s). We present a tool that applies such ideas to the *personal web*, that is, the space rooted in user documents but tightly connected to web documents as well. The tool organizes the personal web with a single topic hierarchy based on d ...

Keywords: context, information navigation, personal web, recommender, reconnaissance

17 Analysis and testing of Web applications
Filippo Ricca, Paolo Tonella
July 2001 **Proceedings of the 23rd International Conference on Software Engineering**
Full text available: pdf(167.58 KB) Additional Information: full citation, abstract, references, citings, index terms

Publisher Site

The economic relevance of Web applications increases the importance of controlling and improving their quality. Moreover, the new available technologies for their development allow the insertion of sophisticated functions, but often leave the developers responsible for their organization and evolution. As a consequence, a high demand is emerging for methodologies and tools for quality assurance of Web based systems.

In this paper, a UML model of Web applications is proposed for their ...

Keywords: UML modeling, code analysis, reverse engineering, testing, web applications

18 Translation of web queries using anchor text mining

Wen-Hsiang Lu, Lee-Feng Chien, Hsi-Jian Lee

June 2002 **ACM Transactions on Asian Language Information Processing (TALIP)**, Volume 1 Issue 2

Full text available: pdf(290.79 KB) Additional Information: full citation, abstract, references, citings, index terms

This article presents an approach to automatically extracting translations of Web query terms through mining of Web anchor texts and link structures. One of the existing difficulties in cross-language information retrieval (CLIR) and Web search is the lack of appropriate translations of new terminology and proper names. The proposed approach successfully exploits the anchor-text resources and reduces the existing difficulties of query term translation. Many query terms that cannot be obtained in ...

Keywords: anchor text mining, comparable corpora, cross-language information retrieval, machine translation, parallel corpora, web mining

19 Performance Workload Char. and Adaptation: Improving web performance by client characterization driven server adaptation

Balachander Krishnamurthy, Craig E. Wills

May 2002 **Proceedings of the 11th international conference on World Wide Web**

Full text available: pdf(241.76 KB) Additional Information: full citation, abstract, references, citings, index terms

We categorize the set of clients communicating with a server on the Web based on information that can be determined by the server. The Web server uses the information to direct tailored actions. Users with poor connectivity may choose not to stay at a Web site if it takes a long time to receive a page, even if the Web server at the site is not the bottleneck. Retaining such clients may be of interest to a Web site. Better connected clients can receive enhanced representations of Web pages, such ...

Keywords: client characterization, client connectivity, server adaptation

20 Crawling the web: Building domain-specific web collections for scientific digital libraries: a meta-search enhanced focused crawling method

Jialun Qin, Yilu Zhou, Michael Chau

June 2004 **Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries**

Full text available: pdf(214.74 KB) Additional Information: full citation, abstract, references, index terms

Collecting domain-specific documents from the Web using focused crawlers has been considered one of the most important strategies to build digital libraries that serve the scientific community. However, because most focused crawlers use local search algorithms to traverse the Web space, they could be easily trapped within a limited sub-graph of the Web that surrounds the starting URLs and build domain-specific collections that are not comprehensive and diverse enough to scientists and researcher ...

Keywords: digital libraries, domain-specific collection building, focused crawling, meta-search, web search algorithm

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next